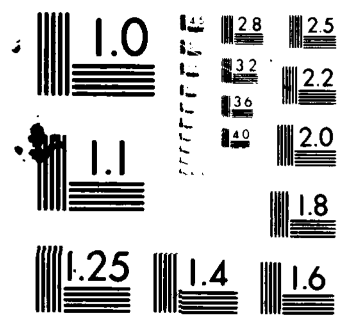
AD-A110 490 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
MEAN VALUE AND TAYLOR FORMS IN INTERVAL ANALYSIS.(U)
OCT 81 L B RALL DAAG29-80-C-0041
UNCLASSIFIED MRC-TSR-2286 NL

1 of 1
AD
A110490

END
DATE
FILMED
3 82
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS

LEVEL II

AD A110490

MRC Technical Summary Report # 2286

MEAN VALUE AND TAYLOR FORMS IN INTERVAL ANALYSIS

L. B. Rall

DTIC ELECTE FEB 4 1982 H

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

October 1981

(Received September 28, 1981)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

MEAN VALUE AND TAYLOR FORMS IN INTERVAL ANALYSIS

L. B. Rall

Technical Summary Report #2286

October 1981

ABSTRACT

Basic spaces for interval analysis are constructed as Cartesian products of the real line. The spaces obtained in this way include real finite and infinite dimensional real vector spaces, and have a number of important Hilbert and Banach spaces as subspaces in the sense of set inclusion. A Gâteaux-type derivative is defined in these spaces, and is used in the corresponding interval spaces, together with interval arithmetic, to obtain interval versions of the mean value theorem and Taylor's theorem. These theorems provide ways to construct accurate interval inclusions of operators, called mean value and Taylor forms. The forms resulting from expansion about midpoints of intervals are shown to be inclusion monotone, and the effect of outward rounding on this class of forms is also considered. An application is made to show that interval iteration operators for the solution of operator equations can be constructed which have arbitrarily high order of convergence in width. Derivations of the fundamental theorems of less generality from results in real and functional analysis are also presented. As in the case of real and functional analysis, the interval Taylor's theorem given here provides a powerful tool for applications of interval analysis to problems in applied mathematics.

AMS (MOS) Subject Classifications: 65J05, 65G10, 47A60, 26B12, 26E15, 26E20, 26E25

Key words and phrases: Interval analysis, Calculus in abstract spaces, Mean value theorem, Taylor's theorem, Interval inclusions, Interval iteration, Fixed point problems, Solution of equations

Work Unit Numbers 1 and 3 (Applied Analysis and Numerical Analysis & Computer Science)

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

SIGNIFICANCE AND EXPLANATION

Interval analysis can be used in applications, for example, to compute an interval which includes the range of responses of a physical system to a range of input conditions. In order for this interval result to be useful, however, it should not be unrealistically larger than the range of responses which would actually be observed, as sometimes happens for straightforward use of simple interval arithmetic. One way to form accurate interval inclusions of real transformations can be based on the mean value theorem and Taylor's theorem of ordinary analysis. In this paper, the basic theory and proofs of such theorems in interval analysis are given. The interval Taylor's theorem, for example, gives computable lower and upper bounds for the truncation error in using the Taylor polynomial in place of the corresponding nonlinear operator. Since this technique is often used in applied mathematics, the theory in this paper permits application of interval analysis to the same types of problems. In particular, it is shown how to obtain interval counterparts of rapidly convergent iteration operators for the solution of equations which inherit the same order of convergence, and thus can be efficient in actual computation.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

MEAN VALUE AND TAYLOR FORMS IN INTERVAL ANALYSIS

L. B. Rall

1. A setting for interval analysis. In the same way that real analysis is concerned with transformations of real numbers (or vectors) into others, *interval analysis* [6], [7] deals with transformations of intervals (or interval vectors). Since an ordering relationship is fundamental to the definition of intervals, a natural abstract setting for interval analysis is a partially ordered space [1], [14], or, more specifically, a lattice [1], [8]. Here, a more concrete approach will be taken, which results in the construction of what will be called *IR-spaces* by forming Cartesian products of the set IR of nonempty closed intervals

(1.1) $$X = [a,b] = \{x \mid a \le x \le b,\ x \in R\},$$

on the real line R. Interval analysis on these IR-spaces will be called *real* interval analysis; it is general enough to cover many important applications, and the theory obtained adapts readily to actual numerical computation, for which only a finite set of real numbers is available.

1.1. Real spaces. The spaces to be considered here are built in a natural way from the set R of real numbers. Given a set A, one can form the *Cartesian product*

(1.2) $$P = \prod_A R$$

of R over the *index set* A to obtain a set of *vectors* f with real *components* $f_\alpha \in R$, $\alpha \in A$. Writing $f = \{f_\alpha \mid \alpha \in A,\ f_\alpha \in R\}$ for $f \in P$, P is a *linear* space for the *componentwise* definitions of *addition* f + g and multiplication by *scalars* (real numbers) $a \cdot f$ given by

(1.3) $$f + g = \{f_\alpha + g_\alpha \mid \alpha \in A\},\quad a \cdot f = \{a \cdot f_\alpha \mid \alpha \in A\},$$

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

respectively [12], [16].

Another way of looking at the product space P given by (1.2) is as the set of all *functionals* (real-valued functions) f on A; one writes $f_\alpha = f(\alpha)$, and (1.3) gives the natural definitions of sums and scalar multiples of functionals, which are also functionals.

Definition 1.1. A *real space* (or *R-space* for short) is a linear space P constructed according to (1.2) and (1.3) or the Cartesian product

$$(1.4) \qquad P = \prod_{\beta \in B} P_\beta$$

of such spaces, again with addition and multiplication by scalars defined componentwise.

Examples of R-spaces abound. The choice $A = \{1,2,\ldots,n\}$ in (1.2) gives $P = R^n$, the space of n-dimensional real vectors $f = (f_1, f_2, \ldots, f_n)$, while $A = \{1,2,3,\ldots\}$, the set of positive integers, gives R^∞, which consists of the real sequential vectors $f = (f_1, f_2, f_3, \ldots)$. Going on to $A = X = [a,b]$, a nonempty interval (1.1), one gets $P = R[a,b]$, the space of all real functions f on the interval [a,b], the components of which are usually denoted by $f(x) = f_x$, $a \leq x \leq b$. Similarly, if $Y = [c,d]$ is also an interval, then taking $A = X \times Y = [a,b] \times [c,d]$ gives the space $R([a,b] \times [c,d])$ of real functions f of two variables with components $f(x,y)$, $a \leq x \leq b$, $c \leq y \leq d$, and so on.

Cartesian products (1.4) can be used for concise description of sets of functions taking on values in R-spaces. For example, with $X = [a,b]$, $Y = [c,d]$, the real space $R(X \times Y) \times R(X \times Y)$ consists of all functions $f: X \times Y \subset R^2 \rightarrow R^2$ with components $f(x,y) = (f_1(x,y), f_2(x,y))$, $a \leq x \leq b$, $c \leq y \leq d$. More generally, if $D \subset P$ and Q is a real space, then

$$(1.5) \qquad \prod_D Q = \{f \mid f: D \subset P \rightarrow Q\}$$

is also a real space by Definition 1.1. In (1.5), it is not required that P be a real space, but this will usually be the case in the following discussion. A simple, but important, example of (1.5) is obtained for $P = R^n$, $Q = R^m$, which gives the set

of functions (or *operators*) $f: D \subset R^n \to R^m$, which are fundamental to computational numerical analysis [10], [12].

The subject of *functional analysis* is concerned with analysis on *normed* linear spaces (usually the ones which are *complete*, called *Banach* spaces) [12], [16]. A number of useful spaces of this type over the real scalar field can be considered to be subspaces of real spaces P in the sense that all their elements belong to P. In particular, all finite-dimensional real normed linear spaces are pretty much indistinguishable, due to the equivalence of norms [16], and can be identified with the real spaces R^n. The situation is different for infinite-dimensional spaces. For example, the elements of the Banach space R_∞^∞ of sequential real vectors f such that

$$\|f\|_\infty = \sup_{(n)}\{|f_n|\} < +\infty, \tag{1.6}$$

form a subspace of R^∞ which is different from the one consisting of elements of R_2^∞, for which

$$\|f\|_2 = \{\sum_{n=1} f_n^2\}^{1/2} < +\infty. \tag{1.7}$$

Similarly, the space $C[a,b]$ of *continuous* functions f on $a \leq x \leq b$ (with the usual norm) can be identified with a subspace of $R[a,b]$ which is different from the one obtained from $L_2[a,b]$, which consists of $f \in R[a,b]$ such that

$$\|f\|_2 = \{(L)\int_a^b f(x)^2 dx\}^{1/2} < +\infty, \tag{1.8}$$

where (L) denotes Lebesgue integration [16]. This natural type of embedding of normed linear spaces into real spaces will be helpful below in connection with the derivation of interval versions of results from real and functional analysis.

1.2. <u>Real interval spaces</u>. The set of finite, nonempty intervals (1.1) on the real line R will be denoted by IR. There is a natural identification of real numbers $x \in R$ with *degenerate* intervals $[x,x] \in IR$ with equal endpoints, and one writes

$$x = [x,x]. \tag{1.9}$$

Ordinary arithmetic, extended from R to IR, is called *interval arithmetic* [6], [7]. For example, *addition* of intervals $X = [a,b]$ and $Y = [c,d]$ is defined by

$$(1.10)\qquad X + Y = [a,b] + [c,d] = [a + c, b + d],$$

and *multiplication* of $X = [a,b]$ by a real number $r = [r,r]$ by

$$(1.11)\qquad r\cdot X = \begin{cases} [ra,rb], & r \geq 0, \\ [rb,ra], & r < 0. \end{cases}$$

Note that with these definitions, IR is *not* a linear space; with subtraction defined in the usual way by $X - Y = X + (-1)\cdot Y$, (1.10) and (1.11) give

$$(1.12)\qquad [0,1] - [0,1] = [-1,1]$$

instead of the identity element $0 = [0,0]$ of interval addition.

It will be useful to associate the following real numbers with an interval $X = [a,b] \in IR$: Its *midpoint* $m(X) = m[a,b] = (a + b)/2$, its *width* $w(X) = w[a,b] = b - a$, and its *absolute value* (or *modulus*) $|X| = |[a,b]| = \max\{|a|,|b|\}$ [7].

Another important property of intervals is that the *intersection* $X \cap Y = [a,b] \cap [c,d]$ is either the interval

$$(1.13)\qquad X \cap Y = [a,b] \cap [c,d] = [\max\{a,c\},\min\{b,d\}]$$

or the *empty set* $\emptyset$; if $b < c$ or $a > d$, then

$$(1.14)\qquad X \cap Y = \emptyset,$$

otherwise, (1.13) holds. Furthermore, if $\{X_n\}$ is a sequence of *nested* intervals, that is

$$(1.15)\qquad X_1 \supset X_2 \supset X_3 \supset \dots ,$$

then

$$(1.16)\qquad X = \bigcap_{n=1}^{\infty} X_n \neq \emptyset,$$

since each X_n is a closed, nonempty subset of R [15].

The construction (1.2), (1.4) of real spaces in §1.1 will now be used to obtain

the corresponding interval spaces, by starting with IR in place of R.

Definition 1.2. A *real interval space* IP is a space of the form

(1.17) $$IP = \prod_{A} IR,$$

or

(1.18) $$IP = \prod_{\beta \in B} IP_{\beta},$$

in which each real interval space IP_{β} is of the form (1.17). Real interval spaces will also be referred to as *IR-spaces*.

There is an obvious one-to-one correspondence between interval spaces (1.17), (1.18) and real spaces (1.2), (1.4), respectively. Furthermore, the order relationships $<$, $\leq$, $\geq$, $>$, in R can be extended componentwise to a real space P to obtain a *partial ordering* [1] of P. In the resulting partial ordering, the corresponding IR-space IP consists of the set of all *intervals* in P; that is, $X \in IP$ if and only there are elements $a,b \in P$ such that $a \leq b$ and $X = [a,b] = \{x \mid a \leq x \leq b,\ x \in P\}$, which is (1.1) with R replaced by P. This leads to the embedding $x = [x,x]$ of P into IP, as in (1.9). Moreover, interval arithmetic is also extended componentwise from IR to an arbitrary real interval space IP. As in the case of IR, IP will not be a linear space, unlike its underlying R-space P. The quantities $m(X)$, $w(X)$, and $|X|$ defined previously for real intervals $X \in IR$ can also be defined componentwise for $X \in IP$, with the result being that $m(X)$, $w(X)$, and $|X|$ will be elements of P.

Typical examples of IR-space are the space IR^n of *interval vectors*

(1.19) $$X = (X_1, X_2, \ldots, X_n), \quad X_i \in IR, \quad i = 1,2,\ldots,n,$$

and the space IR[a,b] of *interval functions* Y on $[a,b] \in IR$ defined by

(1.20) $$Y(x) = [c(x), d(x)], \quad a \leq x \leq b,$$

where $c,d \in R[a,b]$ and $c < d$ [3], [13]. For $X \in IR^n$, for example, one has

(1.21) $$m(X) = (m(X_1), m(X_2), \ldots, m(X_n)) \in R^n,$$

and for $Y \in IR[a,b]$, $|Y|$ is defined by

(1.22) $$|Y|(x) = |Y(x)| = \max\{|c(x)|, |d(x)|\}, \quad a \le x \le b,$$

and thus $|Y| \in R[a,b]$ is a real function.

An interval $X \in IP$ is, by construction, a subset of the underlying R-space P. One important property of intervals in IP as subsets of P is *convexity*.

Lemma 1.1. If P is a real space and $X \in IP$, then X is a *convex* subset of P, that is, for arbitrary points $x,y \in X$,

(1.23) $$\Lambda(x,y) = \{z \mid z = \theta y + (1-\theta)x, \ 0 \le \theta \le 1\} \subset X.$$

Proof: It follows from Definition 1.1 that each $f \in P$, P a real space, can be represented as $f = \{f_\gamma \mid f_\gamma \in R, \ \gamma \in B \times A = \Gamma\}$, the real numbers f_γ, $\gamma \in \Gamma$, being the *components* of f. Now, let $X = [a,b]$, and define $c,d \in P$ by

(1.24) $$c_\gamma = \min\{x_\gamma, y_\gamma\}, \quad d_\gamma = \max\{x_\gamma, y_\gamma\}, \quad \gamma \in \Gamma.$$

For $x,y \in X$, it follows that

(1.25) $$a_\gamma \le c_\gamma \le \theta y_\gamma + (1-\theta)x_\gamma \le d_\gamma \le b_\gamma, \quad \gamma \in \Gamma,$$

for $0 \le \theta \le 1$; hence, from (1.23), $\Lambda(x,y) \subset X$. QED.

As usual, the set $\Lambda(x,y)$ defined by (1.23) is called the *line segment from* x ($\theta = 0$) *to* y ($\theta = 1$). A useful class of intervals are the *symmetric* intervals, defined as follows:

Definition 1.3. An interval $S \in IP$ is said to be *symmetric* if $-s \in S$ for each $s \in S$.

As a consequence of this definition, each symmetric interval S contains the *origin* 0 of P; furthermore, $S = [-a,a]$ for some element $a \ge 0$ of P. Moreover, if $s \in P$, then $S = [-1,1]\cdot s$ will be a symmetric interval; in this case, one can write $S = [-1,1]\cdot s = s\cdot[-1,1] = [-|s|,|s|]$, where $|s|$ is the *absolute value* of s defined componentwise in the usual way.

In addition to the origin 0 of a real space P (the element such that $0_\gamma = 0$,

$\gamma \in \Gamma$), it is helpful to single out the element $e \in P$ defined by $e_\gamma = 1$, $\gamma \in \Gamma$. In terms of e, the symmetric intervals $S_\rho \in IP$ are defined for real $\rho \geq 0$ by

(1.26) $$S_\rho = \rho\cdot[-e,e] = \rho e\cdot[-1,1], \quad \rho \in R, \quad \rho \geq 0.$$

Definition 1.4. A set $D \subset P$ is said to be *bounded* if

(1.27) $$D \subset S_\rho = \rho e\cdot[-1,1]$$

for some real ρ such that $0 \leq \rho < +\infty$, in particular, if D consists of a single element $f \in P$, then f is called a *bounded element* of P.

2. Interval transformations. Suppose that IP, IQ are IR-spaces, and $F: ID \subset IP \rightarrow IQ$ is an operator defined on a domain ID in IP which takes on values in IQ. The result of applying F to $X \in ID$ is symbolized by

(2.1) $$Y = F(X),$$

where $Y \in IQ$, and F is called an *interval transformation* from $ID \subset IP$ into IQ. It follows that $F \in \prod_{ID} IQ$. What will be called *interval analysis* here refers to the study of interval transformations.

Definition 2.1. The interval transformation $F: ID \subset IP \rightarrow IQ$, where IP, IQ are real interval spaces, is said to have an *interval domain* ID if $Z \in ID$ implies that $X \in ID$ for each subinterval $X \subset Z$ of Z.

An important class of interval transformations are the ones which are *monotone* in the sense of the following definition.

Definition 2.2. An interval transformation $F: ID \subset IP \rightarrow IQ$ with interval domain ID is said to be *inclusion monotone* (or simply *monotone*) on ID if

(2.2) $$X \subset Z \Rightarrow F(X) \subset F(Z)$$

for each $Z \in ID$.

Given a domain $D \subset P$, P a real space, the corresponding interval domain ID in IP can be constructed from the set of intervals $Z \subset D$ (which includes all the degenerate intervals equivalent to points of D) by adjoining all subintervals of each

such Z, if necessary. In what follows, it will be assumed that domains ID for interval transformations corresponding to domains D of real transformations are formed in this way, and hence will be interval domains. One has also $D \subset ID$ by the identification of points of P with degenerate intervals in IP. The concept of an interval domain corresponding to a real leads to a fundamental relationship between real and interval transformations.

Definition 2.3. The interval transformation $F: ID \subset IP \to IQ$ is said to be an *inclusion* of the real transformation $f: D \subset P \to Q$ between the underlying real spaces P and Q if

$$(2.3) \qquad f(X) = \{f(x) \mid x \in X\} \subset F(X)$$

for each $X \subset D$. If F is monotone on ID, then it is called a *monotone inclusion* of f.

For most of the results to be obtained below, inclusions of real transformations are adequate. However, the property of monotonicity is highly desirable in many applications. Some interval inclusions of real transformations also have the following property.

Definition 2.4. The interval inclusion $F: ID \subset IP \to IQ$ of $f: D \subset P \to Q$ is said to have the *restriction property* on D if

$$(2.4) \qquad F(x) = F([x,x]) = f(x)$$

for each $x \in D$, in which case F is called an *interval extension* of f on D. If F is also monotone on D, then it is called a *monotone interval extension* of f on D.

The rules of interval arithmetic [6], [7] are examples of monotone interval extensions, in this case of the real transformations $f: R^2 \to R$ defined by $f(x,y) = x \circ y$ for $\circ = +, -, \cdot, /$. (For division, $D = R^2 \setminus \{0\}$, of course.) In actual computation, one ordinarily has to forego the restriction property (2.4), since it is impossible to represent arbitrary real numbers exactly with the finite set of numbers available on a given computer. The use of interval arithmetic and *directed* (or perhaps *outward*) *rounding*, however, allows one to construct monotone inclusions of rational functions automatically, even if the endpoints of intervals have to be selected

from a finite set of numbers G, provided that the computation stays within the interval $IG = [\min\{G\}, \max\{G\}]$ [6], [7]. Along with interval arithmetic, there are other methods for the construction of interval inclusions of real transformations. The ones to be discussed in this paper are based on interval versions of the mean value theorem and Taylor's theorem in ordinary real analysis [4].

3. A derivative in R-spaces. As usual, if P is a real space, then a function $f: D \subset R \to P$ will be called an *abstract function*; for example, $z: R \to P$ defined for $x, y \in P$ by

$$z(\theta) = \theta y + (1-\theta)x = x + \theta(y-x), \quad \theta \in R, \tag{3.1}$$

takes on values on the *line* through x,y for $x \neq y$ (see (1.23)). For $f: D \subset R \to P$, where D contains a neighborhood of 0, it is said that

$$\lim_{\theta \to 0} f(\theta) = 0 \tag{3.2}$$

if there is a real-valued function $\rho \geq 0$, monotone decreasing in $|\theta|$, such that

$$f(\theta) \in \rho(\theta)e \cdot [-1,1] \quad \text{and} \quad \lim_{\theta \to 0} \rho(\theta) = 0. \tag{3.3}$$

Definition 3.1. A function $f: D \subset P \to Q$, D convex, is said to be *differentiable* at $x \in D$ if a linear operator, denoted by $f'_D(x)$ or simply $f'(x)$, exists from the linear space LD spanned by D into Q such that

$$\lim_{\theta \to 0} \frac{r_{x,y}(\theta)}{\theta} = 0, \tag{3.4}$$

where

$$r_{x,y}(\theta) = f(x + \theta(y-x)) - f(x) - f'(x) \cdot \theta(y-x). \tag{3.5}$$

The operator $f'(x)$, easily seen to be unique if it exists, is of course called the *derivative* of f at $x \in D$. (The linear space LD referred to in Definition 3.1 is simply the set of all linear combinations of elements of D [16].) Defined in this way, f' is a derivative of Gâteaux type; in fact, if D is a Banach subspace of P such

that condition (3.4) and $\lim_{\theta\to 0}\|\frac{1}{\theta}\cdot r_{x,y}(\theta)\| = 0$ are equivalent (such as R^n, R_ω^∞, and $C[0,1]$), then $f'(x)$ is precisely the Gâteaux derivative in D of f at x [4], [10]. Because of the dependence of $f'(x) = f_D'(x)$ on the domain D, this derivative can also be considered to be a type of *directional* derivative; for example, one can take D to be the line through the origin of P consisting of the points defined by (3.1). In case it is desirable to distinguish the derivative defined above from some other derivative, it will be called the *elementary* real derivative, or simply the *R-derivative* of f at $x \in D$.

4. <u>Elementary mean value forms.</u>

Theorem 4.1. If X is an interval such that f is differentiable on $X \cap D$, D convex, and F' is an interval inclusion of f' on X, then

$$f(y) - f(x) \in F'(X)\cdot(X - x), \quad x,y \in X \cap D. \tag{4.1}$$

Proof. For $x,y \in X \cap D$, it follows from Definition 3.1 that given a real $\epsilon > 0$, there exists a real number τ, $0 < \tau \leq 1$, such that

$$f(x + \theta(y - x)) - f(x) \in F'(X)\cdot\theta(y - x) + \epsilon\theta e\cdot[-1,1] \tag{4.2}$$

for $0 < \theta < \tau$. To show that (4.2) holds for $\theta = 1$, the assumption that $\tau < 1$ is the supremum of the values for which it is valid will now be contradicted. Set $z = x + \tau(y - x)$. Since $f'(z)$ exists, there is a real number β, $\tau < \beta \leq 1$ such that

$$f(x + \eta(y - x)) - f(z) \in F'(X)\cdot(\eta - \tau)(y - x) + \epsilon(\eta - \tau)e\cdot[-1,1], \tag{4.3}$$

$\tau \leq \eta < \beta$. Let $\theta = \tau$ in (4.2) and add to (4.3) to obtain

$$f(x + \eta(y - x)) - f(x) \in F'(X)\cdot\eta(y - x) + \epsilon\eta e\cdot[-1,1], \tag{4.4}$$

$\tau \leq \eta < \beta$, and thus (4.2) holds for $0 < \theta < \beta$, which contradicts the assumed property of τ, since $\beta > \tau$. Hence, for $\theta = 1$, $\epsilon = 1/n$, (4.2) becomes

$$f(y) - f(x) \in F'(X)\cdot(X - x) + \frac{e}{n}\cdot[-1,1]. \tag{4.5}$$

It follows that

$$f(y) - f(x) \in \bigcap_{n=1}^{\infty}\{F'(X)\cdot(X - x) + \frac{e}{n}\cdot[-1,1]\} = F'(X)\cdot(X - x) + [0,0], \tag{4.6}$$

which is nothing more nor less than (4.1). QED.

The proof of Theorem 4.1 given above is truly elementary in that only interval arithmetic and the definitions of interval inclusions and the derivative are used. Replacing X by $\Lambda(x,y)$ in the above proof leads to the conclusion

$$f(y) - f(x) \in F'(\Lambda(x,y))\cdot(y - x) \subset F'(\Lambda(x,y))\cdot(X - x), \tag{4.7}$$

which is also valid. If F' is a monotone inclusion of f', then (4.7) implies (4.1). Note that f' need not be defined on all of X; all that is required is that $f'(X \cap D) \subset F'(X)$; one can take $f'(x) = F'([x,x])$ for $x \in X\backslash\{X \cap D\}$.

Definition 4.1. If F' is an interval inclusion of f' on X, then the interval inclusion F of f on X defined by

$$F(X) = f(x) + F'(X)\cdot(X - x) \tag{4.8}$$

is called the *(elementary) mean value form* of f.

The mean value form was introduced by Moore [6] in R^n, and studied in R^n and $C'[a,b]$ by Caprani and Madsen [2], whose results will be returned to below. The form (4.8) provides a method, in addition to interval arithmetic, for the construction of interval inclusions of real transformations. A useful case of the mean value form is its *midpoint* (or *centered*) form, obtained for $x = m(X)$. Since

$$X = m(X) + \frac{1}{2}w(x)\cdot[-1,1] \tag{4.9}$$

for an arbitrary interval X and $w(X) \geq 0$, one has, if $m(X) \in D$,

$$F(X) = f(m(X)) + \frac{1}{2}|F'(X)|w(X)\cdot[-1,1] \tag{4.10}$$

in this case, which expresses F(X) as the sum of the point $f(m(X)) \in Q$ and a symmetric interval in IQ. The midpoint form (4.10) is even simpler in case $y \in D$ and

$$X = X(y,\rho) = y + \rho e\cdot[-1,1] \tag{4.11}$$

is the *cube* with center $y \in P$ and radius ρ. Then,

$$F(X(y,\rho)) = f(y) + \rho|F'(X(y,\rho))|e\cdot[-1,1], \tag{4.12}$$

which often can be computed very economically.

The following theorem, which is a generalization of the fundamental result due to Caprani and Madsen [2], shows that F defined by the midpoint mean value form (4.10) is monotone.

Theorem 4.2. If F' is a monotone inclusion of f', then F defined by the midpoint mean value form (4.10) is monotone on the set of intervals X such that $m(X) \in D$.

Since F is already an inclusion of f on the set of intervals cited, all that needs to be established is monotonicity. The following lemma is the key to the proof.

Lemma 4.1 (Caprani-Madsen [2]). If X,Z are intervals in a real space P, then

$$(4.13) \qquad X \subset Z \Leftrightarrow \frac{1}{2}w(Z) \geq \frac{1}{2}w(X) + |m(Z) - m(X)|.$$

Proof: Suppose the inequality in (4.13) holds. Then, for $x \in X$,

$$(4.14) \qquad x - m(Z) = x - m(X) + \{m(X) - m(Z)\} \in \{\frac{1}{2}w(X) + |m(Z) - m(X)|\}\cdot[-1,1],$$

so that $x \in m(Z) + \frac{1}{2}w(Z)\cdot[-1,1] = Z$, and thus $X \subset Z$. On the other hand, suppose that $X \subset Z$, or

$$(4.15) \qquad m(X) + \frac{1}{2}w(X)\cdot[-1,1] \subset m(Z) + \frac{1}{2}w(Z)\cdot[-1,1].$$

Since $w(Z) \geq w(X) \geq 0$, this gives

$$(4.16) \qquad m(X) - m(Z) \in \frac{1}{2}\{w(Z) - w(X)\}\cdot[-1,1],$$

and the inequality in (4.13) follows from multiplication by $[-1,1]$. QED.

Proof of Theorem 4.2. Suppose that $U \subset V$, where $U,V \in IP$ are such that $m(U)$, $m(V) \in D$. Set $X = F(U)$, $Z = F(V)$. It follows that $m(X) = f(m(U))$, $m(Z) = f(m(V))$, $\frac{1}{2}w(X) = |F'(U)|\frac{w(U)}{2}$, $\frac{1}{2}w(Z) = |F'(V)|\frac{w(V)}{2}$. Since $U \subset V$, one has $m(U), m(V) \in V$, and, from the proof of Theorem 4.1,

$$(4.17) \qquad f(m(V)) - f(m(U)) \in F'(V)\{m(V) - m(U)\},$$

so that

$$(4.18) \qquad f(m(V)) - f(m(U)) \in |F'(V)||m(V) - m(U)|\cdot[-1,1].$$

For $x \in X$, $x - m(Z) = x - m(X) + f(m(U)) - f(m(V))$, and

$$(4.19) \qquad x - m(X) \in \frac{1}{2}|F'(U)|w(U)\cdot[-1,1] \subset \frac{1}{2}|F'(V)|w(U)\cdot[-1,1],$$

since the monotonicity of F' implies that $|F'(V)| \geq |F'(U)|$ for $U \subset V$. Using (4.18) and (4.19), one gets

$$x - m(Z) \in |F'(V)|\{\tfrac{1}{2}w(U) + |m(V) - m(U)|\}\cdot[-1,1] \subset \tfrac{1}{2}|F'(V)|w(V)\cdot[-1,1] \tag{4.20}$$

by the Caprani-Madsen Lemma 4.1, so that $x \in Z$, and thus $U \subset V \Rightarrow F(U) \subset F(V)$. QED.

Monotonicity is often crucial in numerical computation, in which only a finite set of points G and corresponding intervals IG are available. When an interval $X \in$ IP is approximated by an interval $Z \in$ IG $\subset$ IP such that $X \subset Z$ (this process is called *outward rounding*), one wants to be sure that $F(X) \subset F(Z)$ in order for the results actually computed to contain the ones that would be obtained by exact computation. In connection with approximate computation, there is also the problem that $f(m(X))$ ordinarily cannot be evaluated exactly. Monotonicity can be preserved in this case on some interval $Y \in$ IG if for each $x \in Y$, there is an element $z(x) \in Q$ which can be computed exactly such that

$$f(x) \in z(x) + \varepsilon e\cdot[-1,1], \quad x \in Y, \tag{4.21}$$

for some known $\varepsilon > 0$. The interval inclusion F of f defined by

$$F(X) = z(m(x)) + \{\tfrac{1}{2}F'(X)w(X) + \varepsilon e\}\cdot[-1,1] \tag{4.22}$$

will then be monotone on *subintervals* of Y for monotone F'; that is, $F(X) \subset F(Z)$ for $X \subset Z \subset Y$. Since actual computation is limited to some interval Y defined by the largest and smallest available real numbers, this type of monotonicity is satisfactory for practical purposes.

5. Elementary Taylor forms. It can be verified without difficulty that the elementary derivative defined in §3 has the ordinary properties of a Gâteaux derivative, for example, $(f + g)' = f' + g'$ and the chain rule holds; proofs will be omitted here. Furthermore, successive differentiations give rise to *multilinear* operators from P into Q in the usual way [4], [10], [12]. The following result is an interval version of Taylor's theorem of real analysis.

Theorem 5.1. If f is differentiable n times on $X \cap D$, D convex, and $F^{(n)}$ is an

interval inclusion of $f^{(n)}$ on X, then for $x,y \in X \cap D$,

$$(5.1) \qquad f(y) - f(x) - \sum_{k=1}^{n-1} \frac{1}{k!} f^{(k)}(x)(y-x)^k \in \frac{1}{n!} F^{(n)}(X)\cdot(X-x)^n.$$

Proof. The proof will be carried out by mathematical induction. Theorem 4.1 shows that (5.1) is valid for $n = 1$, and it will be assumed to hold for $n = m - 1$. If ϕ is an abstract function which is differentiable on $[0,1]$, then, given any $\varepsilon > 0$, it follows as in the proof of Theorem 4.1 that there exists a finite sequence of points $\{\theta_i\}_{i=0}^{\nu}$, $0 = \theta_0 < \theta_1 < \dots < \theta_{\nu-1} < \theta_\nu = 1$, such that

$$(5.2) \qquad \phi(\theta_i) - \phi(\theta_{i-1}) \in \phi'(\theta_{i-1})(\theta_i - \theta_{i-1}) + \varepsilon(\theta_i - \theta_{i-1})e\cdot[-1,1].$$

For the particular abstract function

$$(5.3) \qquad \phi(\theta) = f(x + \theta(y-x)) - f(x) - \sum_{k=1}^{m-1} \frac{1}{k!} f^{(k)}(x)\theta^k(y-x)^k,$$

one has $\phi(0) = 0$, and thus

$$(5.4) \qquad \phi(1) - \phi(0) = \phi(1) = f(y) - f(x) - \sum_{k=1}^{m-1} \frac{1}{k!} f^{(k)}(x)(y-x)^k,$$

and

$$(5.5) \qquad \phi'(\theta) = f'(x + \theta(y-x))(y-x) - f'(x)(y-x) - \sum_{k=2}^{m-1} \frac{1}{(k-1)!} f^{(k)}(x)\theta^{k-1}(y-x)^k.$$

By the induction hypothesis,

$$(5.6) \qquad \phi'(\theta) \in \frac{1}{(m-1)!} F^{(m)}(X)\cdot(X-x)^m\theta^{m-1}[0,1].$$

Therefore, from (5.2),

$$(5.7) \qquad \phi(\theta_i) - \phi(\theta_{i-1}) \in \frac{1}{(m-1)!} F^{(m)}(X)\cdot(X-x)^m\theta_{i-1}^{m-1}(\theta_i - \theta_{i-1})[0,1] + \\ + \varepsilon(\theta_i - \theta_{i-1})e\cdot[-1,1],$$

$i = 1,2,\dots,\nu$. Thus,

$$(5.8) \qquad \phi(1) - \phi(0) = \sum_{i=1}^{\nu} \{\phi(\theta_i) - \phi(\theta_{i-1})\} \in \frac{1}{(m-1)!} F^{(m)}(X)(X-x)^m \sum_{i=1}^{\nu} \theta_{i-1}^{m-1}(\theta_i - \theta_{i-1})\cdot[0,1] + \\ + \varepsilon e\cdot[-1,1].$$

However,

$$(5.9) \qquad 0 < \sum_{i=1}^{\nu} \theta_{i-1}^{m-1}(\theta_i - \theta_{i-1}) < \int_0^1 \theta^{m-1} d\theta = \frac{1}{m},$$

since the sum is a lower Riemann sum for the integral. Since X is convex and $x \in X$ implies $0 \in (X - x)$, it follows that $(X - x)\cdot\alpha[0,1] = \alpha(X - x) \subset (X - x)\cdot\beta[0,1] = \beta(X - x)$ for $0 < \alpha < \beta$. Using this fact and

$$(5.10) \qquad \bigcap_{\varepsilon\to 0} \varepsilon e\cdot[-1,1] = [0,0],$$

one has

$$(5.11) \qquad \phi(1) \in \frac{1}{m!} F^{(m)}(X)\cdot(X - x)^m + [0,0] = \frac{1}{m!} F^{(m)}(X)\cdot(X - x)^m,$$

which is equivalent to (5.1) with $n = m$ by (5.4). This completes the proof of the theorem by mathematical induction. QED.

Once again, little more than interval arithmetic is required in the proof.

Definition 5.1. If $f: D \subset P \to Q$ is differentiable n times on $X \cap D$, $X \in IP$, then for $x \in X \cap D$,

$$(5.12) \qquad F(X) = f(x) + \sum_{k=1}^{n-1} \frac{1}{k!} f^{(k)}(x)\cdot(X - x)^k + \frac{1}{n!} F^{(n)}(X)\cdot(X - x)^n,$$

where $F^{(n)}$ is an interval inclusion of $f^{(n)}$ on X, is called the *(elementary) Taylor form* of f of *order* n.

It follows from Theorem 5.1 that F defined by (5.12) is an interval inclusion of f on X. For the particular choice $x = m(X)$, one obtains the *midpoint* form of (5.12),

$$(5.13) \qquad F(X) = f(m(X)) + \{\sum_{k=1}^{n-1} \frac{1}{2^k k!}|f^{(k)}(m(X))|w(X)^k + \frac{1}{2^n n!}|F^{(n)}(X)|w(X)^n\}\cdot[-1,1],$$

and, for $X = X(y,\rho)$ a cube, the *cube-centered* form

$$(5.14) \qquad F(X(y,\rho)) = f(y) + \{\sum_{k=1}^{n-1} \frac{\rho^k}{k!}|f^{(k)}(y)|e^k + \frac{\rho^n}{n!}|F^{(n)}(X(y,\rho))|e^n\}\cdot[-1,1].$$

Evaluations of this latter form can often be carried out very economically, since

operations on e ordinarily do not require multiplications, and only non-negative operators are involved. Monotonicity of the midpoint Taylor form (5.13) also follows from monotonicity of $F^{(n)}$, in much the same way as for the midpoint mean value form (4.10).

Theorem 5.2. If $F^{(n)}$ is a monotone inclusion of $f^{(n)}$, then F defined by the midpoint Taylor form (5.13) is monotone on the set of intervals X such that $m(X) \in D$.

Proof. As before, suppose $U \subset V$, and it is to be shown that $F(U) \subset F(V)$, where the results of the transformations of U,V by F are given by (5.13). For brevity of notation, set $u = m(U)$, $v = m(V)$. It follows that

$$(5.15) \qquad m(F(U)) = f(m(U)) = f(u), \quad m(F(V)) = f(m(V)) = f(v),$$

and

$$(5.16) \qquad \tfrac{1}{2}w(F(U)) = |f'(u)|\tfrac{1}{2}w(U) + \frac{1}{2!}|f''(u)|(\tfrac{1}{2}w(U))^2 + \ldots + \frac{1}{n!}|F^{(n)}(U)|(\tfrac{1}{2}w(U))^n,$$

with an analogous expression for $\frac{1}{2}w(F(V))$. In order to prove that $F(U) \subset F(V)$, it will be shown that $\frac{1}{2}w(F(U)) + |m(F(U)) - m(F(V))| \le \frac{1}{2}w(F(V))$, from which the desired result follows by the Caprani-Madsen Lemma 4.1.

First, since $F^{(n)}$ is assumed to be monotone,

$$(5.17) \qquad |F^{(n)}(U)|(\tfrac{1}{2}w(U))^n \le |F^{(n)}(V)|(\tfrac{1}{2}w(U))^n.$$

Furthermore, by Theorem 5.1,

$$(5.18) \qquad |f^{(k)}(u)| \le |f^{(k)}(v)| + \sum_{j=k+1}^{n-1} \frac{1}{(j-k)!}|f^{(j)}(v)|\cdot|u-v|^{j-k} + \frac{1}{(n-k)!}|F^{(n)}(V)||u-v|^{n-k}$$

$$= \sum_{j=k}^{n-1} \frac{1}{(j-k)!}|f^{(j)}(v)|\cdot|u-v|^{j-k} + \frac{1}{(n-k)!}|F^{(n)}(V)|\cdot|u-v|^{n-k},$$

$k = 1,2,\ldots,n-1$, using the result of multiplication of (5.1) by $[-1,1]$. It follows from (5.16), (5.17), and (5.18) that

$$(5.19) \qquad \tfrac{1}{2}w(F(U)) \le \sum_{k=1}^{n-1} \frac{1}{k!}\Big(\sum_{j=k}^{n-1} \frac{1}{(j-k)!}|f^{(j)}(v)||u-v|^{j-k}\Big)(\tfrac{1}{2}w(U))^k$$

$$+ \sum_{k=1}^{n} \frac{1}{k!(n-k)!}|F^{(n)}(V)||u-v|^{n-k}(\tfrac{1}{2}w(U))^k.$$

Interchange of order of the double summation in (5.19) results in

$$(5.20)\qquad \sum_{k=1}^{n-1}\sum_{j=k}^{n-1}\frac{1}{k!(j-k)!}|f^{(j)}(v)|\cdot|u-v|^{j-k}(\tfrac{1}{2}w(U))^k$$

$$= \sum_{j=1}^{n-1}\sum_{k=1}^{j}\frac{1}{k!(j-k)!}|f^{(j)}(v)|\cdot|u-v|^{j-k}(\tfrac{1}{2}w(U))^k.$$

Theorem 5.1 also gives

$$(5.21)\qquad |f(u) - f(v)| \le \sum_{j=1}^{n-1}\frac{1}{j!}|f^{(j)}(v)|\cdot|u-v|^j + \frac{1}{n!}|F^{(n)}(V)|\cdot|u-v|^n.$$

Addition of (5.19) and (5.21) results in the inequality

$$(5.22)\qquad \tfrac{1}{2}w(F(U)) + |f(u) - f(v)| \le \sum_{j=1}^{n-1}\frac{1}{j!}|f^{(j)}(v)|\sum_{k=0}^{j}\frac{j!}{k!(j-k)!}|u-v|^{j-k}(\tfrac{1}{2}w(U))^k$$

$$+ \frac{1}{n!}|F^{(n)}(V)|\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}|u-v|^{n-k}(\tfrac{1}{2}w(U))^k$$

$$= \sum_{j=1}^{n-1}\frac{1}{j!}|f^{(j)}(v)|\{\tfrac{1}{2}w(U) + |u-v|\}^j$$

$$+ \frac{1}{n!}|F^{(n)}(V)|\{\tfrac{1}{2}w(U) + |u-v|\}^n.$$

Hence, by the Caprani-Madsen Lemma 4.1,

$$(5.23)\qquad \tfrac{1}{2}w(F(U)) + |m(F(U)) - m(F(V))| \le \sum_{j=1}^{n-1}\frac{1}{j!}|f^{(j)}(v)|(\tfrac{1}{2}w(V))^j + \frac{1}{n!}|F^{(n)}(V)|(\tfrac{1}{2}w(V))^n$$

$$= \tfrac{1}{2}w(F(V)),$$

and thus $F(U) \subset F(V)$. QED.

Remark 5.1. Some of the combinatorial aspects of the proof of Theorem 5.2 can be avoided by the use of the identity

$$(5.24)\qquad \phi(u) + \phi'(u)(x-u) + \ldots + \frac{1}{(n-1)!}\phi^{(n-1)}(u)(x-u)^{n-1}$$

$$= \phi(v) + \phi'(v)\{(x-u) + (u-v)\} + \ldots + \frac{1}{(n-1)!}\phi^{(n-1)}(v)\{(x-u) + (u-v)\}^{n-1},$$

in which ϕ is an (abstract) polynomial of degree n-1 [12].

As in the case of the midpoint mean value form (4.10), the corresponding Taylor form (5.13) requires modification in actual computation. Suppose that for some interval $Y \in IG$, there is for each $x \in Y$ a point $z(x) \in Q$ and operators $z_j(x): Y^j \to Q$, $j = 1,2,\ldots,n-1$, which can be computed exactly, such that

$$(5.25) \qquad f(x) \in z(x) + \varepsilon e \cdot [-1,1], \quad f^{(j)}(x) \in z_j(x) + \varepsilon_j \cdot [-1,1],$$

where $\varepsilon \geq 0$ is a real number and $\varepsilon_j \geq 0$ for each $\varepsilon_j: Y^j \to Q$, $i = 1,2,\ldots,n-1$. Then, the *rounded Taylor operator* F defined by

$$(5.26) \qquad F(X) = z(x) + \{\varepsilon e + \sum_{j=1}^{n-[illegible]} [illegible] | + \varepsilon_j](\tfrac{1}{2}w(X))^j + \frac{1}{n!}|F^{(n)}(X)|(\tfrac{1}{2}w(X))^n\} \cdot [-1,1]$$

is inclusion monotone on intervals $X \subseteq Y$, which will usually be satisfactory for use in actual computation.

6. <u>Application to iteration operators</u>. The interval versions of the mean value and Taylor's theorem given above, like their counterparts in real and functional analysis, have numerous applications. Theorem 5.1 shows, for example, that the *interval remainder term*

$$(6.1) \qquad R_n(X) = \frac{1}{n!}F^{(n)}(X) \cdot (X - x)^n$$

contains the *truncation error* $f(y) - f_{n-1}(y)$ resulting from the use of the *Taylor polynomial*

$$(6.2) \qquad f_{n-1}(y) = f(x) + f'(x)(y - x) + \ldots + \frac{1}{(n-1)!}f^{(n-1)}(x)(y - x)^{n-1}$$

of degree $n - 1$ in place of $f(y)$ for arbitrary $y \in X$. In particular, the results obtained by Moore [6], [7] on the numerical solution of differential equations by interval methods follow from this expansion.

The application to be considered here is to the solution of the equation

$$(6.3) \qquad f(x) = 0$$

for $x \in D \to P$, where $f: D \subset P \to Q$ is a differentiable operator. Given a nonsingular linear operator $Y: Q \to P$, equation (6.3) can be transformed into the *fixed point problem* $x = g(x)$ for the operator $g: D \subset P \to P$ defined by

(6.4) $$g(x) = x - Yf(x).$$

Since simple iteration is often used to solve fixed point problems, g will be called an *iteration operator* for f. The choice $Y = f'(x)^{-1}$ corresponds to *Newton's method* for the solution of (6.3), $Y = f'(z)^{-1}$ for $z \neq x$ to a method of *Stirling type* [11], and so on. Treating Y as a constant operator, one has

(6.5) $$g'(x) = I - Yf'(x),\ g'(x) = -Yf''(x),\ \ldots,\ g^{(n)}(x) = -Yf^{(n)}(x),$$

where I denotes the identity operator in P, and thus, if f is differentiable at least n times, then

(6.6) $$g(x) \in z - Yf(z) + \{I - Yf'(z)\}(x - z) - \ldots - \frac{1}{(n-1)!}f^{(n-1)}(z)(x - z)^{n-1} - \frac{1}{n!}YF^{(n)}(X)\cdot(X - z)^n$$

for $x,z \in X$, where $F^{(n)}$ is an interval inclusion of $f^{(n)}$ on X.

Now, given an arbitrary sequence $Y_0, Y_1, \ldots$ of nonsingular linear operators, a sequence of intervals $X_0, X_1, \ldots,$ and points $z_k \in X_k$, $k = 0,1,2,\ldots,$ one can construct the corresponding sequence $G_0, G_1, \ldots$ of *interval iteration operators* for f defined by

(6.7) $$G_k(X_k) = z_k - Yf(z_k) + \{I - Y_k f'(z_k)\}\cdot(X_k - z_k) - \ldots - \frac{1}{(n-1)!}Y_k f^{(n-1)}(z_k)\cdot(X_k - z_k)^{n-1} - \frac{1}{n!}YF^{(n)}(X_k)\cdot(X_k - z_k)^n,$$

$k = 0,1,2,\ldots$. The following theorem is a direct consequence of (6.6).

Theorem 6.1. If $x = x^* \in X_0$ is a solution of (6.3), then for

(6.8) $$X_{k+1} = X_k \cap G_k(X_k),\quad k = 0,1,2,\ldots,$$

one has

(6.9) $$x^* \in X = \bigcap_{k=0}^{\infty} X_k.$$

Proof: It follows from (6.6) that $x^* \in X_k \Rightarrow x^* \in G_k(X_k)$, since $x^* = g(x^*)$, which

in turn implies $x^* \in X_{k+1}$. This gives (6.9). QED.

The process (6.9) is called *interval iteration* [14]. Since

$$X_0 \supset X_1 \supset X_2 \supset \ldots, \tag{6.10}$$

this process gives improved lower and upper bounds for x^* as long as $X_{k+1} \neq X_k$. (Of course, if $X_{k+1} = X_k$, then $X = X_k$, and the interval iteration terminates in a finite number of steps.) The contrapositive of the assertion in Theorem 6.1 is that if

$$X_{k+1} = X_k \cap G_k(X_k) = \emptyset \tag{6.11}$$

for some positive integer k, where $\emptyset$ denotes the empty set, then $x^* \notin X_0$, and there is consequently no fixed point of g or solution of (6.3) in the initial interval X_0 [14].

In the case n = 1, one obtains the *Krawczyk operators* [5]

$$K_k(X_k) = z_k - Y_k f(z_k) + \{I - Y_k F'(X_k)\} \cdot (X_k - z_k) \tag{6.12}$$

from (6.7), with $z_k = m(X_k)$. Suppose that F'' is an interval inclusion of f'' which is *consistent* with F' in the sense that

$$L - F'(X) \subset F''(X)w(X), \quad L \in F'(X). \tag{6.13}$$

Then, from (6.12), for $Y_k^{-1} \in F'(X_k)$,

$$K_k(X_k) \subset z_k - Y_k f(z_k) + \frac{1}{2} Y_k F''(X_k) w(X_k)^2 \cdot [-1,1], \tag{6.14}$$

since $X_k - z_k = X_k - m(X_k) = \frac{1}{2} w(X_k) \cdot [-1,1]$. It follows that interval iteration with the Krawczyk operator converges quadratically as $w(X_k) \to 0$ to a degenerate interval, thus mimicking the behavior of its real counterparts.

For n = 2, the Chebyshev-type iteration operator [12]

$$T_k(X_k) = z_k - Y_k f(z_k) + \{I - Y_k f'(z_k)\} \cdot (X_k - z_k) - \frac{1}{2} Y_k F''(X_k) \cdot (X_k - z_k)^2 \tag{6.15}$$

results, and so on. It follows that (6.7) can be used to construct interval iteration operators with arbitrarily high orders of convergence in width as $w(X_k) \to 0$.

7. <u>*Other derivations of the mean value and Taylor's theorems.*</u> *In certain* particular cases, Taylor's theorem as given above (which includes the mean value theorem for $n = 1$), can be derived directly from classical results in real or functional analysis. For example, with $P = Q = R$, one has

$$(7.1) \qquad f(b) = f(a) + f'(a)(b - a) + \ldots + \frac{1}{(n-1)!}f^{(n-1)}(a)(b - a)^{n-1} + \frac{1}{n!}f^{(n)}(\xi)(b - a)^n, \quad a < \xi < b,$$

in which the remainder term is said to be in *Lagrange* form. For $X = [a,b]$, one has $f^{(n)}(\xi) \in F^{(n)}(X)$, $b - a \in X - a$, which gives (5.1) at once in this special case. Formula (7.1) also hold componentwise in R^ν, which leads to a similar generalization, since $f_k^{(n)}(\xi_k) \in F_k^{(n)}(X)$, $k = 1,2,\ldots,\nu$, even though (7.1) does not necessarily hold for some $\xi \in X \subset R^\nu$. This generalization to R^ν has been used by Moore [6], [7], and Caprani and Madsen [2]. In the latter paper, a version of the mean value theorem was also derived for integral operators, but the results are not easy to interpret without the use of interval integration [3], [13].

A more straightforward method of generalization of Taylor's theorem can be based on the use of the *Cauchy* form of the remainder term,

$$(7.2) \qquad R_n(f;a,b) = \int_0^1 f^{(n)}(a + \theta(b - a))\frac{(1 - \theta)^{n-1}}{(n-1)!}\cdot(b - a)^n d\theta,$$

which holds in Banach spaces [4], [12]. In R, the use of interval integration gives

$$(7.3) \qquad R_n(f;a,b) \in \int_0^1 F^{(n)}(a + \theta(b-a))\frac{(1-\theta)^{n-1}}{(n-1)!}(b-a)^n d\theta \subset F^{(n)}(X)\cdot(X-a)^n\int_0^1 \frac{(1-\theta)^{n-1}}{(n-1)!}d\theta,$$

from which (5.1) is obtained by evaluation of the real integral. By a simple extension of the concept of the interval integral [3], [13] to abstract functions f which take on values in a Banach space D, $D \subset Q$, Q a real space, a corresponding generalization of formula (7.3) will be obtained.

In order to construct the interval integral of an abstract function $\phi = \phi(\theta)$ which takes on values in an R-space Q for $0 \le \theta \le 1$, one simply partitions $\Theta = [0,1]$

into subintervals $\Theta_i = [\theta_{i-1}, \theta_i]$, $i = 1,2,\ldots,m$, by means of points $0 = \theta_0 \le \theta_1 \le \ldots \le \theta_m = 1$. The set of all such partitions into m subintervals will be denoted by Δ_m. The *abstract interval function* Φ: $I\Theta \to IQ$ will be defined by

$$(7.4) \qquad \Phi(\Theta_i) = [\inf_{\Theta_i} \phi(\theta), \sup_{\Theta_i} \phi(\theta)].$$

Definition 7.1. The *interval integral* of the abstract function ϕ over $[0,1]$ is defined to be

$$(7.5) \qquad \int_0^1 \phi(\theta)d\theta = \bigcap_{m=1}^{\infty} \bigcap_{\Delta_m} \sum_{i=1}^{m} \Phi(\Theta_i)w(\Theta_i) \in IQ.$$

This follows exactly the construction of [3]; again, the interval integral defined by (7.5) is the intersection of a nested sequence of nonempty intervals, and hence is nonempty.

Now, suppose that $D \subset Q$ is a Banach space in which $X \cap D$ is a closed set for $X \in IQ$. The Riemann (R) integral of abstract functions ϕ taking on values in D is defined to be the limit of the *Riemann sums*

$$(7.6) \qquad \Sigma_{m,\Delta}\phi = \sum_{i=1}^{m} \phi(\tau_i)(\theta_i - \theta_{i-1}), \quad \tau_i \in \Theta_i,$$

as $m \to \infty$ and $\|\Delta\| = \max_{(i)} w(\Theta_i) \to 0$ [4], [12]. It follows that

$$(7.7) \qquad (R)\int_0^1 \phi(\theta)d\theta \in \sum_{i=1}^{m} \Phi(\Theta_i)w(\Theta_i) \subset \Phi(\Theta),$$

since the intersection of D with the *interval Darboux sums* [3] appearing in (7.5) is closed in the topology of D. Therefore, from (7.5),

$$(7.8) \qquad (R)\int_0^1 \phi(\theta)d\theta \in \int_0^1 \phi(\theta)d\theta,$$

if ϕ is Riemann (R) integrable over [0,1] in the sense defined by Graves [4]. Thus, in the special case that f is a function taking on values in a Banach space D with $f^{(n)}(a + \theta(b - a))$ Riemann integrable over [0,1], (7.3) follows immediately by interval integration, and gives (5.1) for interval inclusions $F^{(n)}$ of $f^{(n)}$. This derivation is also less general than the one given in §5, which holds in R-spaces.

References

1. G. Birkhoff: Lattice Theory, Amer. Math. Soc. Colloq. Pubs., Vol. 25, Rev. Ed., American Mathematical Society, New York, 1948.

2. O. Caprani and K. Madsen: Mean value forms in interval analysis, Computing 25 (1980), 147-154.

3. O. Caprani, K. Madsen, and L. B. Rall: Integration of interval functions, SIAM J. Math. Anal. 12 (1981), 321-341.

4. L. M. Graves: Riemann integration and Taylor's theorem in general analysis, Trans. Amer. Math. Soc. 29 (1927), 163-177.

5. R. Krawczyk: Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken, Computing 4 (1969), 187-201.

6. R. E. Moore: Interval Analysis, Prentice-Hall, Englewood Cliffs, N. J., 1966.

7. R. E. Moore: Methods and Applications of Interval Analysis, SIAM Studies in Applied Mathematics, 2, Society for Industrial and Applied Mathematics, Philadelphia, 1979.

8. K. Nickel: Verbandstheoretische Grundlagen der Intervall-Mathematik, [9], pp. 251-262 (1975).

9. K. Nickel (Ed.): Interval Mathematics, Lecture Notes in Computer Science No. 29, Springer-Verlag, Berlin-Heidelberg-New York, 1975.

10. J. M. Ortega and W. C. Rheinboldt: Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

11. L. B. Rall: Convergence of Stirling's method in Banach spaces, Aequationes Math. 12 (1975), 12-20.

12. L. B. Rall: Computational Solution of Nonlinear Operator Equations, Wiley, New York, 1969; reprinted by Krieger, Huntington, N. Y., 1979.

13. L. B. Rall: Integration of interval functions II. The finite case, SIAM J. Math. Anal. (to appear)

14. L. B. Rall: A theory of interval iteration, MRC Tech. Summary Rept. No. 2196, University of Wisconsin-Madison, 1981.

15. W. Sierpiński: General Topology, Tr. by C. Cecilia Krieger, University of Toronto Press, Toronto, 1952.

16. A. E. Taylor: Introduction to Functional Analysis, Wiley, New York, 1958.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TSR# 2286	2. GOVT ACCESSION NO. AD-A110 490	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MEAN VALUE AND TAYLOR FORMS IN INTERVAL ANALYSIS		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) L. B. Rall		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 1 and 3 (Applied Analysis & Numerical Analysis & Computer Science)
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE October 1981
		13. NUMBER OF PAGES 23
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Interval analysis, Calculus in abstract spaces, Mean value theorem, Taylor's theorem, Interval inclusions, Interval iteration, Fixed point problems, Solution of equations

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Basic spaces for interval analysis are constructed as Cartesian products of the real line. The spaces obtained in this way include real finite and infinite dimensional real vector spaces, and have a number of important Hilbert and Banach spaces as subspaces in the sense of set inclusion. A Gâteaux-type derivative is defined in these spaces, and is used in the corresponding interval spaces, together with interval arithmetic, to obtain interval versions of the mean value theorem and Taylor's theorem. These theorems provide ways to construct accurate

20. **ABSTRACT, cont.**

interval inclusions of operators, called mean value and Taylor forms. The forms resulting from expansion about midpoints of intervals are shown to be inclusion monotine, and the effect of outward rounding on this class of forms is also considered. An application is made to show that interval iteration operators for the solution of operator equations can be constructed which have arbitrarily high order of convergence in width. Derivations of the fundamental theorems of less generality from results in real and functional analysis are also presented. As in the case of real and functional analysis, the interval Taylor's theorem given here provides a powerful tool for applications of interval analysis to problems in applied mathematics.